A STATISTICAL FRAMEWORK TO CHARACTERISE MICROSTRUCTURE IN HIGH ANGULAR RESOLUTION DIFFUSION IMAGING

Sofia Olhede and Brandon Whitcher

Department of Statistical Science, University College London, London, UK,
Clinical Imaging Centre, GlaxoSmithKline, London, UK.

ABSTRACT

High angular resolution diffusion imaging (HARDI) is an MR acquisition technique that has been used to identify and estimate multi-modal structure in white matter. We introduce a multiresolution estimator of the diffusion density in q-space based on a wavelet lifting scheme for HARDI data. The proposed technique avoids imposing isotropic smoothing with a fixed bandwidth that is a common feature in other higher-order methods. Summary statistics for the properties of the diffusion density are derived entirely in q-space, based on great circle and perpendicular great circle statistics. A framework is developed to characterise structural features of the diffusion using these statistics.

Index Terms— Anisotropy, HARDI measurements, smoothing, wavelets.

1. INTRODUCTION

HARDI measurements of the diffusion in q-space are collected in a set of orientations $\{\tilde{\mathbf{q}}_i\}_{i=1}^n$ on the unit sphere. The diffusion probability density function (PDF), denoted $a(\mathbf{x})$, is in the absence of measurement error linked to the normalized measured signal $A(\mathbf{q})$ via an inverse Fourier transform. HARDI sampling permits the determination of orientational inhomogeneity of the PDF, e.g. using multi-tensor methods, q-ball imaging or PAS-MRI [1, 2].

Multi-tensor modelling performs well if the mixture population is indeed Gaussian and the number of fibre populations is known *a priori* [3]. Both q-ball imaging and PAS-MRI require smoothing for relatively consistent estimation of the diffusion PDF. This means that high-frequency features are removed over an interval of frequencies, regardless of their contribution to the underlying structure. PAS-MRI also has a smoothness parameter that must be fixed.

We develop a q-space density estimator with *variable bandwidth* based on constructing wavelet coefficients directly from the fixed gradient sampling scheme on the unit sphere. The q-space density is estimated by applying a universal thresholding rule to the set of derived wavelet coefficients. The wavelet transform is inverted to estimate the q-space density. Methods using a single conservative threshold are prone to oversmoothing, where high-frequency coefficients associated with structure rather than noise are needlessly removed. To alleviate such problems Cai and Silverman [4] proposed thresholding a given coefficient based on the value of its neighbouring coefficients, a rule that we adapt to the neighbourhood defined by the gradient encoding scheme. This multiresolution method reduces the artificial symmetry imposed on reconstructions by other non-parametric estimation methods.

Once a density estimator has been calculated in q-space, we propose to characterise its properties directly in q-space rather than applying additional transformations into image space. This is in contrast to previous methods; e.g., using the Funk-Radon Transform (FRT) to produce an orientation distribution function (ODF) [1]. One problem with existing methods of calculating ODFs is that their interpretation varies across methods [3], and they cannot be directly interpreted as PDFs in space.

Interpretable summaries must be computed from a non-parametric density estimator. Scalar measures in space from the FRT were defined by [1]. We introduce the great circle and perpendicular great circle statistics that form the backbone for characterising complicated spatial properties of the diffusion signal in q-space. This characterisation is achieved in a series of steps by determining unidirectionality, degree of anisotropy and symmetry in space, directly from the q-space summaries.

2. ESTIMATING THE Q-SPACE DENSITY

Given a set of HARDI measurements in q-space $\{\tilde{\mathbf{q}}_i\}_{i=1}^n$, assume that the observed variables $A(\tilde{\mathbf{q}}_i)$ have a Rician distribution with intensity $\mathcal{A}(\tilde{\mathbf{q}}_i)$ and error variance σ^2. Estimation methods for $\mathcal{A}(\tilde{\mathbf{q}}_i)$ are commonly

The authors thank EPSRC and GlaxoSmithKline for supporting the research via EP/E031536/1.

built on assuming a degree of smoothness of $\mathcal{A}(\mathbf{q})$, or using a parametric model. We use *lifting* [5] to construct wavelet coefficients $\{\widehat{W}_{ji}\}$ from the observed data, where the index j denotes scale and i denotes the location. The transform is constructed from distances that are based on great circles, using the Haar transform [5]. The nearest coefficient to location i at any scale j is denoted $b_{ji} \neq i \neq b_{j'i},\ \forall j' < j$. We use the redundant wavelet transform to avoid misalignment problems from local shifts between the underlying density and the sampling grid. We terminate the transform at the *primary resolution* J, and complete the representation with the empirical scaling coefficients $\{\widehat{V}_{Ji}\}$. With the empirical wavelet coefficients, it is possible to estimate the variance of the noise process and derive bias-corrected estimators of the scaling coefficients using the Rician distributional assumptions. The bias-corrected scaling coefficients are denoted $\{\widehat{V}_{Ji}^{(\mathrm{db})}\}$.

Threshold estimators for the wavelet coefficients are obtained from a hard thresholding rule

$$\begin{aligned} \widehat{W}_{ji}^{(\mathrm{hts})} &= \begin{cases} \widehat{W}_{ji} & \text{if} \quad |\widehat{W}_{ji}| > \widehat{\sigma}_{ji}\sqrt{2\log(n)}; \\ 0 & \text{if} \quad |\widehat{W}_{ji}| \leq \widehat{\sigma}_{ji}\sqrt{2\log(n)}; \end{cases} \qquad (1) \\ \widehat{\sigma}_{ji}^2 &= 2\widehat{\sigma}^2 + \frac{\widehat{V}_{ji}^{(\mathrm{db})2}}{2^j} - \frac{\pi\widehat{\sigma}^2}{2} L_{1/2}^2\left(-\frac{\widehat{V}_{ji}^{(\mathrm{db})2}}{2^{j+1}\widehat{\sigma}^2}\right), \end{aligned}$$

where $L_{1/2}(x) = e^{x/2}[(1-x)I_0(-x/2) - xI_1(-x/2)]$, with $I_n(\cdot)$ a modified Bessel function of the first kind. The parameter $\widehat{\sigma}$ is estimated from the level $j = 1$ wavelet coefficients using the maximum absolute deviation (MAD) estimator. It is known that universal hard thresholding is inappropriate for a variety of signal-to-noise scenarios. We therefore adopt a local threshold that depends on a subset of wavelet coefficients. We fix (j,i) and define a set of neighbouring coefficients (except at $j = 1, J$) as $\mathcal{S}_{ji} = \{(j,i), (j+1,i), (j,b_{1i}), (j,b_{2i})\}$. The estimated local magnitude is then $\widehat{|W_{ji}|}^2 = \sum_{(j',i')\in\mathcal{S}_{ji}} |\widehat{W}_{j'i'}|^2/S$. If the local variance is slowly varying, and if there is no signal present in S_{ji} we may note that (approximately)

$$\sum_{(j',i')\in\mathcal{S}_{ji}} \left|\widehat{W}_{j'i'}\right|^2 \sim \sigma_{ji}^2 \chi_S^2, \qquad (2)$$

where S is the number of coefficients in $\mathcal{S}_{ji}$. Thus

$$\widehat{W}_{ji}^{(\mathrm{htm})} = \begin{cases} \widehat{W}_{ji} & \text{if} \quad \widehat{|W_{ji}|}^2 > S\lambda_{ji}^2; \\ 0 & \text{if} \quad \widehat{|W_{ji}|}^2 \leq S\lambda_{ji}^2, \end{cases} \qquad (3)$$

where $\lambda_{ji}^2 = \widehat{\sigma}_{ji}^2[2\log(n) + (S-2)\log\log(n)]$. The distributional results are not exact because there may be correlation between coefficients. The appropriate threshold for a χ_S^2 variable is discussed in [6]. Inverting the wavelet transform gives us two non-parametric q-space estimators $\{\widehat{\mathcal{A}}^{(\mathrm{hts})}(\mathbf{q}_i)\}$ and $\{\widehat{\mathcal{A}}^{(\mathrm{htm})}(\mathbf{q}_i)\}$.

3. DETERMINING SPATIAL FEATURES FROM THE Q-SPACE DENSITY

Intuition in q-space, along with suitable summary statistics, are required in order to interpret the estimated density. Firstly, we note that a simple uni-directional structure in image space corresponds to a great circle in q-space [1]. A simple model of a unimodal diffusion is with $B(\cdot)$ a decreasing function, $\boldsymbol{\upsilon}_i \in \mathbb{R}^3$, and $\xi_i > 0$

$$\mathcal{A}(\mathbf{q}) = B\left(\mathbf{q}^T\left[\sum_i \xi_i \boldsymbol{\upsilon}_i \boldsymbol{\upsilon}_i^T\right]\mathbf{q}\right). \qquad (4)$$

If the voxel exhibits purely isotropic diffusion then $\xi_1 = \xi_2 = \xi_3$ and in expectation all great-circles have equal magnitude. To obtain a statistic with adequate power, we start by estimating $\boldsymbol{\upsilon}_1$ by the maximiser of the FRT.

We denote the great circle in q-space, associated with point $\widehat{\boldsymbol{\upsilon}}_1$ in image space, via $\mathcal{G}(\widehat{\boldsymbol{\upsilon}}_1)$. The great circle orthogonal to $\mathbf{q} \in \mathcal{G}(\widehat{\boldsymbol{\upsilon}}_1)$ is $\mathcal{G}^{\perp}(\widehat{\boldsymbol{\upsilon}}_1, \mathbf{q})$. If the density is isotropic, then there is no difference in $\mathcal{A}(\mathbf{q}')$ for $\mathbf{q}' \in \mathcal{G}(\widehat{\boldsymbol{\upsilon}}_1)$ or for $\mathbf{q}' \in \mathcal{G}^{\perp}(\widehat{\boldsymbol{\upsilon}}_1, \mathbf{q})$ for any $\mathbf{q}$. Perceived maxima are then due to random fluctuations alone. In contrast if the density should be modelled by (4) then $\mathcal{A}(\mathbf{q}')$ is constant for $\mathbf{q}' \in \mathcal{G}(\boldsymbol{\upsilon}_1)$ and exhibits decay away from the maximum great circle.

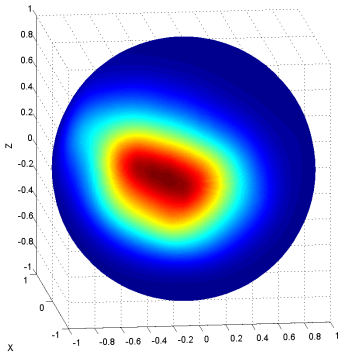
We interpolate $\left\{\widehat{\mathcal{A}}(\tilde{\mathbf{q}}_k)\right\}$ to $\{\hat{\mathbf{q}}_k\}_{k=1}^N$, $\hat{\mathbf{q}}_k \in \mathcal{G}(\widehat{\boldsymbol{\upsilon}}_1)$ for a fixed N, and also the N perpendicular great circles $\mathcal{G}^{\perp}(\widehat{\boldsymbol{\upsilon}}_1, \hat{\mathbf{q}}_k)$ using a Delaunay triangulation and triangle-based linear interpolation. We write $\hat{\mathbf{p}}_{jk} = \alpha_j \widehat{\boldsymbol{\upsilon}}_1 \pm \sqrt{1-\alpha_j^2}\hat{\mathbf{q}}_k$. We compute (using interpolation)

$$\widehat{\mathcal{A}}_{\perp}(\alpha_j) = \frac{1}{N}\sum_{k=1}^{N} \widehat{\mathcal{A}}\left(\hat{\mathbf{p}}_{jk}\right). \qquad (5)$$

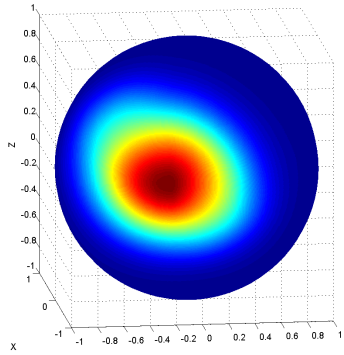
The estimators $\widehat{j}_{\min}$ and $\widehat{j}_{\max}$ are taken to minimise/maximise $\widehat{\mathcal{A}}_{\perp}(\alpha_j)$, respectively, while $\widehat{k}_{\min}$ and $\widehat{k}_{\max}$ are chosen to minimise/maximise $\widehat{\mathcal{A}}(\hat{\mathbf{q}}_k)$. We test $H_0 : \mathcal{A}(\mathbf{q}') = \mathcal{A}$ for all $\mathbf{q}'$ on the great circle and the perpendicular great circles (isotropic diffusion) versus $H_1 : \mathcal{A}(\mathbf{q}')$ is more variable on the perpendicular (unimodal diffusion), corresponding to (4). The test statistic for isotropy is then given by

$$T_m = \left[\frac{\widehat{\mathcal{A}}_{\perp}(\alpha_{\widehat{j}_{\max}})}{\widehat{\mathcal{A}}_{\perp}(\alpha_{\widehat{j}_{\min}})}\right]\left[\frac{\widehat{\mathcal{A}}(\hat{\mathbf{q}}_{\widehat{k}_{\max}})}{\widehat{\mathcal{A}}(\hat{\mathbf{q}}_{\widehat{k}_{\min}})}\right]^{-1} - 1. \qquad (6)$$

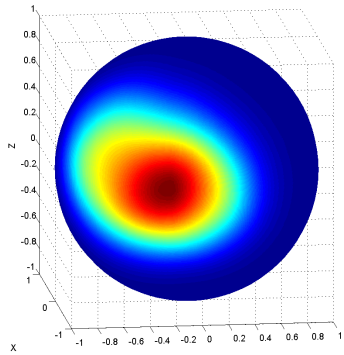
We assume the SNR is sufficiently large so that we may perform a Taylor expansion of $\widehat{\mathcal{A}}(\hat{\mathbf{q}}_k)$ about $\mathcal{A}(\mathbf{q}_k)$ and



(a) The noisy ODF calculated by the FRT.



(b) Thresholding coefficients jointly ($J = 4$).



(c) Thresholding coefficients individually ($J = 4$).

Fig. 1. The estimated unidirectional diffusion.

approximate the estimated distribution as Gaussian with variance $\sigma_{\mathcal{A}}^2$. With $\sigma_{\mathcal{A}}$ and $\mathcal{A}$ known, under the null hypothesis of isotropy where m is the effective degrees of freedom of a great-circle (this depends on the spherical acquisition), we can derive a suitable threshold from maxima of Gaussian random variables, and denote this λ_m. If the statistic is rejected to the right (i.e., $\mathcal{A}T_m/\sigma_{\mathcal{A}} > \lambda_m$) then there is reason to suspect the voxel microstructure is uni-directional, rather than isotropic.

We estimate $\mathcal{A}$ and $\sigma_{\mathcal{A}}$ under H_0 using $\widehat{\mathcal{A}} = \sum_{k=1}^{N} \widehat{\mathcal{A}}(\hat{\mathbf{q}}_k)/N$ and $\widehat{\sigma}_{\mathcal{A}} = \sqrt{3}\text{median}\{|\widehat{\mathcal{A}}(\hat{\mathbf{q}}_k) - \widehat{\mathcal{A}}|\}/0.6745$, where $\hat{\mathbf{q}}_k$ are evenly spaced vectors on the great circle. We note that even if the $\widehat{\mathcal{A}}(\hat{\mathbf{q}}_k)$ are contaminated, the median is robust [7]. We let $U_m = \widehat{\mathcal{A}}T_m/\widehat{\sigma}_{\mathcal{A}}$. This statistic will enable us to test, at each voxel, whether the distribution appears directional, *without* the distribution of the test statistic depending on any unknown nuisance parameters. Some caution must be exercised due to *multiple comparisons* and for example False Discovery Rate could be adapted when testing across all voxels in a large region.

If the hypothesis of isotropy has been rejected the degree of anisotropy is estimated via

$$\widehat{\xi} = \frac{\log \widehat{\mathcal{A}}_{\perp}(0)}{\log \widehat{\mathcal{A}}_{\perp}(1)}. \tag{7}$$

The quantity $\{A(\mathbf{q}_k)\}$ must be normalized in order to estimate ξ [2], and we use the mean of the observations at $\mathbf{q} = 0$. $\widehat{\xi}$ for a Gaussian PDF estimates the ratio of the largest eigenvalue to the "average" of the two lesser eigenvalues. We can investigate the hypothesis of ellipsoidality by looking at $\zeta = \max_k \log[\widehat{\mathcal{A}}(\hat{\mathbf{q}}_k)]/\log[\widehat{\mathcal{A}}(\hat{\mathbf{q}}_{k+N/4})]$, across k. We may also quantify persistent asymmetric structure. A local discrete estimator of asymmetry is calculated via

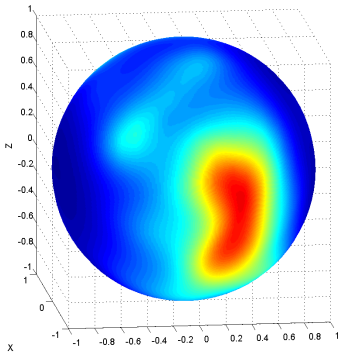
$$\tilde{\kappa}_k = \frac{8\sum_{j=1}^{N/4-1}\left[\widehat{\mathcal{A}}(\hat{\mathbf{p}}_{jk}) - \widehat{\mathcal{A}}(\hat{\mathbf{p}}_{(j+N/4)k})\right]}{\sum_{j=1}^{N}\widehat{\mathcal{A}}(\hat{\mathbf{p}}_{jk})}, \tag{8}$$

$$\breve{k}_{\max} = \arg_k \max \tilde{\kappa}_k,\ \hat{\kappa} = \frac{1}{N/4+1}\sum_{k=\breve{k}_{\max}-N/8}^{\breve{k}_{\max}+N/8}\tilde{\kappa}_k.$$

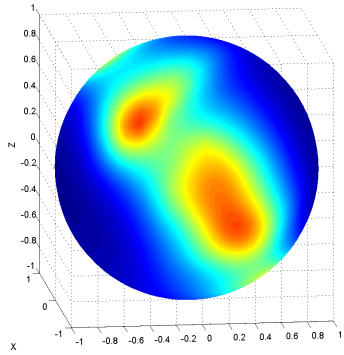
This is approximately (when appropriately scaled) distributed like the absolute value of a Gaussan random variable. If $\hat{\kappa}$ is sufficiently large, then this indicates that $a(\mathbf{x})$ will not be decaying symmetrically.

4. DISCUSSION

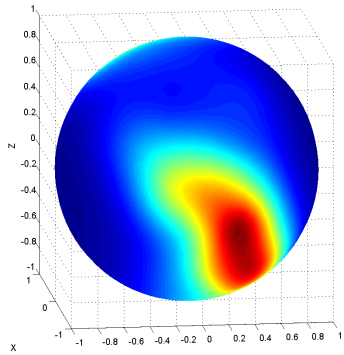
We focus here on determining white matter microstructure from a previously analysed clinical acquisition [8]. We estimate the ODF for a voxel in the splenium of the corpus collosum in Figure 1. The raw FRT displays highly asymmetric structure. The diffusion density estimated by applying the FRT to $\{\widehat{\mathcal{A}}^{(\text{hts})}(\mathbf{q}_i)\}$ is an improvement, but $\{\widehat{\mathcal{A}}^{(\text{htm})}(\mathbf{q}_i)\}$ is visually more appealing. The assumption of isotropy is firmly rejected in favour of uni-modality at this voxel, $U_m = 4.90$ while $\lambda_m = 1.96$ for $\{\widehat{\mathcal{A}}^{(\text{htm})}(\mathbf{q}_i)\}$ ($U_m = 4.90$ for $\{\widehat{\mathcal{A}}^{(\text{hts})}(\mathbf{q}_i)\}$



(a) The rescaled FRT.



(b) Thresholding coefficients jointly ($J = 6$).



(c) Thresholding coefficients individually ($J = 6$).

Fig. 2. The estimated unidirectional diffusion at the superior longitudinal fasiculus.

also, and for the raw measurements $U_m = 0.39 < 1.96$). $\widehat{\xi} = 0.53 < 1$ for $\{\widehat{\mathcal{A}}^{(\text{htm})}(\mathbf{q}_i)\}$ ($\widehat{\xi} = 0.56$ for $\{\widehat{\mathcal{A}}^{(\text{hts})}(\mathbf{q}_i)\}$ and $\widehat{\xi} = 0.59$ for the raw measurements), but none of the other tests are rejected at this voxel. We estimate the density at the intersection of the corticospinal tract and the superior longitudinal fasiculus (Figure 2). The joint thresholding estimator provides a more pleasing reconstruction, without imposing symmetry. For this voxel isotropy is *not* rejected in favour of unimodality; $U_m = -1.09 < 1.96$ for $\{\widehat{\mathcal{A}}^{(\text{htm})}(\mathbf{q}_i)\}$, while $U_m = -0.66 < 1.96$ for $\{\widehat{\mathcal{A}}^{(\text{hts})}(\mathbf{q}_i)\}$, and $U_m = -0.63$ for the raw measurements). The value of $\widehat{\xi}$ reflects that the voxel is a multi-modal and not an isotropic diffusion: $\widehat{\xi} = 0.74$, 0.74 or 0.40 (raw). The other test statistics are not interpretable in this case.

The main goal of this work is to characterise properties of the diffusion density that cannot be captured with existing parametric models or methods that rely on isotropic smoothing. Previous work has focused on investigating the multi-modal nature of diffusion at the voxel level. In contrast, we have characterised the diffusion non-parametrically in terms of its symmetry and directionality. Alternative summaries of the key characteristics of diffusion are a first step towards improving tractography techniques [3].

5. REFERENCES

[1] D. S. Tuch, "Q-ball imaging," *Magnetic Resonance in Medicine*, vol. 52, pp. 1358–1372, 2004.

[2] D. C. Alexander, "Multiple-fibre reconstruction algorithms for diffusion MRI," *Annals of the New York Academy of Sciences*, vol. 1046, pp. 113–133, 2005.

[3] K. Seunarine, P. A. Cook, M. G. Hall, K. V. Embleton, G. J. M. Parker, and D. C. Alexander, "Exploiting peak anisotropy for tracking through complex structure," in *Mathematical Methods in Biomedical Image Analysis*. IEEE Computer Society, Rio de Janeiro, 2007.

[4] T. T. Cai and B. Silverman, "Incorporating information on neighboring coefficients into wavelet estimation," *Sankhya Series B*, vol. 63, pp. 127–148, 2001.

[5] P. Schröder and W. Sweldens, "Spherical wavelets: Efficiently representing functions on the sphere," *Computer Graphics Proceedings (SIGGRAPH 95)*, pp. 161–172, 1995.

[6] T. R. Downie and B. W. Silverman, "The discrete multiple wavelet transform and thresholding methods," *IEEE Transactions on Signal Processing*, pp. 2558–2561, 1998.

[7] G. W. Bassett Jr, "Equivariant, monotonic, 50% breakdown estimators," *The American Statistician*, vol. 45, pp. 135–137, 1991.

[8] B. Whitcher, D. S. Tuch, J. J. Wisco, A. G. Sorensen, and L. Wang, "Using the wild bootstrap to quantify uncertainty in diffusion tensor imaging," Human Brain Mapping, *doi: 10.1002/hbm.20395*, vol. 29, pp. 346–362, 2008.